

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: UNIVERSITE CATHOLIQUE DE LOUVAIN
Halles Universitaires
(B) STREET: Place de l' Universite, 1
(C) CITY: LOUVAIN-LA-NEUVE
(E) COUNTRY: BELGIUM
(F) POSTAL CODE (ZIP): B-1348

(A) NAME: UNIVERSITE DE MONS-HAINAUT
(B) STREET: Place du Parc 20
(C) CITY: MONS
(E) COUNTRY: BELGIUM
(F) POSTAL CODE (ZIP): B-7000

(ii) TITLE OF INVENTION: PEROXISOME-ASSOCIATED PEPTIDE, NUCLEOTIDE
SEQUENCE ENCODING SAID PEPTIDE AND THEIR USES IN THE
DIAGNOSTIC AND/OR THE TREATMENT OF LUNG INJURIES AND
DISEASES, AND OF OXIDATIVE STRESS-RELATED DISORDERS

(iii) NUMBER OF SEQUENCES: 19

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 805 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 193..681

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GCCAGGAGGC GGAGTGAAG TGGCCGTGGG GCGGGTATGG GACTAGCTGG CGTGTGCGCC	60
CTGAGACGCT CAGCGGGCTA TATACTCGTC GGTGGGGCCG GCGGTCAGTC TGCGGCAGCG	120

GCAGCAAGAC GGTGCAGTGA AGGAGAGTGG GCGTCTGGCG GGGTCCGCAG TTTCAGCAGA	180
GCCGCTGCAG CC ATG GCC CCA ATC AAG GTG GGA GAT GCC ATC CCA GCA	228
Met Ala Pro Ile Lys Val Gly Asp Ala Ile Pro Ala	
1 5 10	
GTG GAG GTG TTT GAA GGG GAG CCA GGG AAC AAG GTG AAC CTG GCA GAG	276
Val Glu Val Phe Glu Gly Glu Pro Gly Asn Lys Val Asn Leu Ala Glu	
15 20 25	
CTG TTC AAG GGC AAG AAG GGT GTG CTG TTT GGA GTT CCT GGG GCC TTC	324
Leu Phe Lys Gly Lys Lys Gly Val Leu Phe Gly Val Pro Gly Ala Phe	
30 35 40	
ACC CCT GGA TGT TCC AAG ACA CAC CTG CCA GGG TTT GTG GAG CAG GCT	372
Thr Pro Gly Cys Ser Lys Thr His Leu Pro Gly Phe Val Glu Gln Ala	
45 50 55 60	
GAG GCT CTG AAG GCC AAG GGA GTC CAG GTG GTG GCC TGT CTG AGT GTT	420
Glu Ala Leu Lys Ala Lys Gly Val Gln Val Val Ala Cys Leu Ser Val	
65 70 75	
AAT GAT GCC TTT GTG ACT GGC GAG TGG GGC CGA GCC CAC AAG GCG GAA	468
Asn Asp Ala Phe Val Thr Gly Glu Trp Gly Arg Ala His Lys Ala Glu	
80 85 90	
GGC AAG GTT CGG CTC CTG GCT GAT CCC ACT GGG GCC TTT GGG AAG GAG	516
Gly Lys Val Arg Leu Leu Ala Asp Pro Thr Gly Ala Phe Gly Lys Glu	
95 100 105	
ACA GAC TTA TTA CTA GAT GAT TCG CTG GTG TCC ATC TTT GGG AAT CGA	564
Thr Asp Leu Leu Leu Asp Asp Ser Leu Val Ser Ile Phe Gly Asn Arg	
110 115 120	
CGT CTC AAG AGG TTC TCC ATG GTG GTA CAG GAT GGC ATA GTG AAG GCC	612
Arg Leu Lys Arg Phe Ser Met Val Val Gln Asp Gly Ile Val Lys Ala	
125 130 135 140	
CTG AAT GTG GAA CCA GAT GGC ACA GGC CTC ACC TGC AGC CTG GCA CCC	660
Leu Asn Val Glu Pro Asp Gly Thr Gly Leu Thr Cys Ser Leu Ala Pro	
145 150 155	
AAT ATC ATC TCA CAG CTC TGA GGCCCTGGGC CAGATTACTT CCTCCACCCC	711
Asn Ile Ile Ser Gln Leu *	
160	
TCCCTATCTC ACCTGCCCAG CCCTGTGCTG GGGCCCTGCA ATTGGAATGT TGGCCAGATT	771
TCTGCAATAA ACACTTGTGG TTTGCGGAAA AAAA	805

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 163 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ala Pro Ile Lys Val Gly Asp Ala Ile Pro Ala Val Glu Val Phe
 1 5 10 15
 Glu Gly Glu Pro Gly Asn Lys Val Asn Leu Ala Glu Leu Phe Lys Gly
 20 25 30
 Lys Lys Gly Val Leu Phe Gly Val Pro Gly Ala Phe Thr Pro Gly Cys
 35 40 45
 Ser Lys Thr His Leu Pro Gly Phe Val Glu Gln Ala Glu Ala Leu Lys
 50 55 60
 Ala Lys Gly Val Gln Val Val Ala Cys Leu Ser Val Asn Asp Ala Phe
 65 70 75 80
 Val Thr Gly Glu Trp Gly Arg Ala His Lys Ala Glu Gly Lys Val Arg
 85 90 95
 Leu Leu Ala Asp Pro Thr Gly Ala Phe Gly Lys Glu Thr Asp Leu Leu
 100 105 110
 Leu Asp Asp Ser Leu Val Ser Ile Phe Gly Asn Arg Arg Leu Lys Arg
 115 120 125
 Phe Ser Met Val Val Gln Asp Gly Ile Val Lys Ala Leu Asn Val Glu
 130 135 140
 Pro Asp Gly Thr Gly Leu Thr Cys Ser Leu Ala Pro Asn Ile Ile Ser
 145 150 155 160
 Gln Leu *

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 780 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Rattus Rattus

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 136..624

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TGCCTCCTAG GCAGCATAGC CGGATCGGTG CTCCTGTCAT CGGCTACTTG GACGTGCGTG 60
 GCAGGCAGAG CAGGCCGGA AGGAGCAGGT TGGGAGTGTG GTGGGGCCCG CAGCTTCAGC 120

AGTGCCGCGG TGACTATGGC CCCGATCAAG GTGGGAGACA CCATTCCCTC AGTGGAGGTA	180
TTTGRAGGGG AACCTGGAAG GAAGGTGAAC TTGGCAGAGC TGTTCAAGGA CAAGAAAGGT	240
GTTTTGTTTG GAGTCCCTGG GGCATTTACA CCTGGCTGTT CCAAGACCCA TCTGCCTGGG	300
TTTGTGGAGC AAGCCGGAGC TCYGAAGGCC AAGGGAGCAC AAGTGGTGGC CTGTCTGAGT	360
GTAAATGATG YCTTCGTGAC TGCAGAGTGG GGTTCGAGCCC ACCAGGCAGA AGGCAAGGTT	420
CAGCTCCTGG CTGACCCAC TGGAGCTTTT GGAAAGGAGA CAGATTTACT ACTAGATGAT	480
TCTTTGGTGT CTCTCTTTGG GAATCGTCGG CTAAAAAGGT TCTCCATGGT GATAGACAAG	540
GGCGTAGTAA AGGCACTGAA CGTGGAGCCG GATGGCACAG GCCTCACCTG CAGCCTGGCC	600
CCCAACATCC TCTCACAAC CTGAGGCCCT GACCAGAATG TCCTCTGACT CTCCCATCTC	660
CTCCACCCAG CTCTGGGCCA AAGGCCAGT ACCTCCTTAC CTGAGGGCCA CTGGAATGGA	720
ACCTTGACAA TATTTCTGCA ATAAACAGTT TAATTTGTGA AAAAAAAAAA AAAAAAAAAA	780

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 162 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Rattus Rattus

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION:17
- (D) OTHER INFORMATION:/product= "Glu or Gly"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION:63
- (D) OTHER INFORMATION:/product= "Leu or Pro"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION:79
- (D) OTHER INFORMATION:/product= "Ala or Val"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met	Ala	Pro	Ile	Lys	Val	Gly	Asp	Thr	Ile	Pro	Ser	Val	Glu	Val	Phe
1				5					10					15	

Xaa	Gly	Glu	Pro	Gly	Lys	Lys	Val	Asn	Leu	Ala	Glu	Leu	Phe	Lys	Asp	
			20					25					30			
Lys	Lys	Gly	Val	Leu	Phe	Gly	Val	Pro	Gly	Ala	Phe	Thr	Pro	Gly	Cys	
		35					40					45				
Ser	Lys	Thr	His	Leu	Pro	Gly	Phe	Val	Glu	Gln	Ala	Gly	Ala	Xaa	Lys	
	50					55					60					
Ala	Lys	Gly	Ala	Gln	Val	Val	Ala	Cys	Leu	Ser	Val	Asn	Asp	Xaa	Phe	
65				70						75					80	
Val	Thr	Ala	Glu	Trp	Gly	Arg	Ala	His	Gln	Ala	Glu	Gly	Lys	Val	Gln	
				85					90					95		
Leu	Leu	Ala	Asp	Pro	Thr	Gly	Ala	Phe	Gly	Lys	Glu	Thr	Asp	Leu	Leu	
			100					105					110			
Leu	Asp	Asp	Ser	Leu	Val	Ser	Leu	Phe	Gly	Asn	Arg	Arg	Leu	Lys	Arg	
	115						120					125				
Phe	Ser	Met	Val	Ile	Asp	Lys	Gly	Val	Val	Lys	Ala	Leu	Asn	Val	Glu	
	130					135					140					
Pro	Asp	Gly	Thr	Gly	Leu	Thr	Cys	Ser	Leu	Ala	Pro	Asn	Ile	Leu	Ser	
145					150					155					160	
Gln	Leu															

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 675 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mouse

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 99..588

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TGCTCCGTGC ATCGACGTGC TTGGCAGGCA GAGCAGGCCG GAAAGAAGCA GGTGGGGAGT	60
GTGGCGGAGC CCGCAGCTTC AGCAGCTCCG CGGTGACCAT GGCCCCGATC AAGGTGGGAG	120
ATGCCATTCC CTCAGTGGAG GTATTTGAAG GGGGAACCGG AAAGAAGGTG AACTTGGCAG	180
AGCTGTTCAA GGGCAAGAAA GGTGTTTTGT TTGGAGTCCC TGGGGCATTT ACACCTGGCT	240

GTTCTAAGAC CCACCTGCCT GGGTTTGTGG AGCAAGCTGG AGCTCTGAAG GCTAAGGGAG 300
 CGCAGGTGGT GGCCTGTCTG AGCGTTAATG ACGTCTTTGT GATTGAAGAG TGGGGTCGAG 360
 CCCACCAGGC AGAAGGCAAG GTTCGGCTCC TGGCTGACCC CACTGGAGCC TTTGGGAAGG 420
 CGACAGACTT ATTATTGGAT GATTCTTTGG TGTCTCTCTT TGGGAATCGT CGGCTGAAAA 480
 GGTTCCTCCAT GGTGATAGAC AACGGCATAG TGAAGGCACT GAACGTGGAG CCAGATGGCA 540
 CAGGCCTCAC CTGCAGCCTG GCCCCAACA TCCTCTCCCA ACTCTGAGGC CCTGGCCAGA 600
 TGTCTCTGA CTCTCCCATC TCTCCCACCC GGCTCTAGGC CAAAAGGCTC GGTACCTCCT 660
 TACTGGGAGC CACGT 675

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 162 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Mouse

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met	Ala	Pro	Ile	Lys	Val	Gly	Asp	Ala	Ile	Pro	Ser	Val	Glu	Val	Phe
1				5					10					15	
Glu	Gly	Glu	Pro	Gly	Lys	Lys	Val	Asn	Leu	Ala	Glu	Leu	Phe	Lys	Gly
			20					25					30		
Lys	Lys	Gly	Val	Leu	Phe	Gly	Val	Pro	Gly	Ala	Phe	Thr	Pro	Gly	Cys
		35					40					45			
Ser	Lys	Thr	His	Leu	Pro	Gly	Phe	Val	Glu	Gln	Ala	Gly	Ala	Leu	Lys
	50					55					60				
Ala	Lys	Gly	Ala	Gln	Val	Val	Ala	Cys	Leu	Ser	Val	Asn	Asp	Val	Phe
65					70					75				80	
Val	Ile	Glu	Glu	Trp	Gly	Arg	Ala	His	Gln	Ala	Glu	Gly	Lys	Val	Arg
			85					90						95	
Leu	Leu	Ala	Asp	Pro	Thr	Gly	Ala	Phe	Gly	Lys	Ala	Thr	Asp	Leu	Leu
			100					105					110		
Leu	Asp	Asp	Ser	Leu	Val	Ser	Leu	Phe	Gly	Asn	Arg	Arg	Leu	Lys	Arg
			115				120						125		

Phe Ser Met Val Ile Asp Asn Gly Ile Val Lys Ala Leu Asn Val Glu
130 135 140

Pro Asp Gly Thr Gly Leu Thr Cys Ser Leu Ala Pro Asn Ile Leu Ser
145 150 155 160

Gln Leu

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 469 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 161..382

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGGTATGGGA CTAGCTGGCG TGTGCGCCCT GAGACGCTCA GCGGGCTATA TACTCGTCGG	60
TGGGGCCGGC GGTCAGTCTG CGGCAGCGGC AGCAAGACGG TGCAGTGAAG GAGAGTGGGC	120
GTCTGGCGGG GTCCGCAGTT TCAGCAGAGC CGCTGCAGCC ATGGCCCCAA TCAAGGTTTCG	180
GCTCCTGGCT GATCCCACTG GGGCCTTTGG GAAGGAGACA GACTTATTAC TAGATGATTC	240
GCTGGTGTCC ATCTTTGGGA ATCGACGTCT CAAGAGGTTC TCCATGGTGG TACAGGATGG	300
CATAGTGAAG GCCCTGAATG TGGAACCAGA TGGCACAGGC CTCACCTGCA GCCTGGCACC	360
CAATATCATC TCACAGCTCT GAGGCCCTGG GCCAGATTAC TTCCTCCACC CCTCCCTATC	420
TCACCTGCCC AGCCGTGTGC TGGGGCCCTG CAATTGGAAT GTTGGCCAG	469

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 601 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:161..514

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGGTATGGGA CTAGCTGGCG TGTGCGCCCT GAGACGCTCA GCGGGCTATA TACTCGTCGG	60
TGGGGCCGGC GGTCAGTCTG CGGCAGCGGC AGCAAGACGG TGCAGTGAAG GAGAGTGGGC	120
GTCTGGCGGG GTCCGCAGTT TCAGCAGAGC CGCTGCAGCC ATGGCCCCAA TCAAGACACA	180
CCTGCCAGGG TTTGTGGAGC AGGCTGAGGC TCTGAAGGCC AAGGGAGTCC AGGTGGTGGC	240
CTGTCTGAGT GTTAATGATG CCTTTGTGAC TGGCGAGTGG GGCCGAGCCC ACAAGGCGGA	300
AGGCAAGGTT CGGCTCCTGG CTGATCCCAC TGGGGCCTTT GGGAAGGAGA CAGACTTATT	360
ACTAGATGAT TCGCTGGTGT CCATCTTTGG GAATCGACGT CTCAAGAGGT TCTCCATGGT	420
GGTACAGGAT GGCATAGTGA AGGCCCTGAA TGTGGAACCA GATGGCACAG GCCTCACCTG	480
CAGCCTGGCA CCCAATATCA TCTCACAGCT CTGAGGCCCT GGGCCAGATT ACTTCCTCCA	540
CCCCTCCCTA TCTCACCTGC CCAGCCCTGT GCTGGGGCCC TGCAATTGGA ATGTTGGCCA	600
G	601

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 604 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:161..517

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGGTATGGGA CTAGCTGGCG TGTGCGCCCT GAGACGCTCA GCGGGCTATA TACTCGTCGG	60
---	----

TGGGGCCGGC GGTCACTCTG CGGCAGCGGC AGCAAGACGG TGCAGTGAAG GAGAGTGGGC	120
GTCTGGCGGG GTCCGCAGTT TCAGCAGAGC CGCTGCAGCC ATGGCCCCAA TCAAGGTGGG	180
AGATGCCATC CCAGCAGTGG AGGTGTTTGA AGGGGAGCCA GGAACAAGG TGAACCTGGC	240
AGAGCTGTTC AAGGGCAAGA AGGGTGTGCT GTTTGGAGTT CCTGGGGCCT TCACCCCTGG	300
ATGTTCCAAG GTTCGGCTCC TGGCTGATCC CACTGGGGCC TTTGGGAAGG AGACAGACTT	360
ATTACTAGAT GATTGCTGG TGTCCATCTT TGGGAATCGA CGTCTCAAGA GGTCTCCAT	420
GGTGGTACAG GATGGCATAG TGAAGGCCCT GAATGTGGAA CCAGATGGCA CAGGCCTCAC	480
CTGCAGCCTG GCACCCAATA TCATCTCACA GCTCTGAGGC CCTGGGCCAG ATTACTTCCT	540
CCACCCCTCC CTATCTCACC TGCCCAGCCC TGTGCTGGGG CCCTGCAATT GGAATGTTGG	600
CCAG	604

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2710 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION:2516..2710

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION:2074..2135

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION:1932..1970

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION:1728..1859

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION:802..936

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

TCTGTCCCTT	AGCGCCCCCG	CGGGGGCTTA	CCCCATCCCA	CTCCATGACC	TCCCCTCCCC	60
CCATGGCGAA	TTCCCACCTT	TCTGTCTTTC	ACTCACTTCC	TGGAACCGTC	CCCAGGGCCT	120
TGGACCTTCC	CCCTTCTCCT	CCCAAACCTT	GTGAGACCCC	ATTCCCTTTC	TACTTCATCC	180
TGCTCTCAAC	TTTTGGGCTC	CTCAGAGGCC	CTCACCCCTG	ACTCTCTCTC	CCTACCCACT	240
CTGGTCCCAT	GAAGCCCTCA	AGTACTCTGG	GGATGGATCC	TTCCCCCTTC	AAAAGATTCC	300
TTCTTTTGTT	CTACACCTCC	TGGGTGTAGG	GGCCTGGACA	CCCTCCCCCA	ACGTTCCACC	360
TGCCGCTGCC	CTTCCTCTTC	CTCCTCCTGA	GGGTGGGACC	CTCAGACCTG	GCCAAGATCC	420
TCTCCCTCCA	TGTTGTCAGG	GACTCCTCCT	CACCCCCAAA	TACAGCCCTC	TAGCCCCCTGT	480
CCATTTTATT	CCACTCCTTT	CCTGTAACCT	AGACAGCATG	TTATGCAACC	CTTTGCGACA	540
CATGGGGAAA	CCTTCCCTCC	CTTCCTCTGT	TGTCACCAAT	GGCCCCTTAA	GAGGAGCAGG	600
GCCACCTTGA	AACTTGGAGG	ATATGGGGTA	ACCCAGTGGG	AGCGGGCAGG	GAGGGCCCTT	660
GGAAACTGAC	AGGGCTGGAG	TATCCTGCTG	GGTTTCAGCC	CCGGTTCCTG	CAGGCACAGC	720
TGCCAGGCTC	TCTGTTCAAC	TTCTTGCCTC	TGGTTTGCCC	CGGCTCCCTC	ACCCCCCTTA	780
CCCTGGAGTC	CTTCCTTCTA	GGTGGGAGAT	GCCATCCCAG	CAGTGGAGGT	GTTTGAAGGG	840
GAGCCAGGGA	ACAAGGTGAA	CCTGGCAGAG	CTGTTCAAGG	GCAAGAAGGG	TGTGCTGTTT	900
GGAGTTCCTG	GGGCCTTCAC	CCCTGGATGT	TCCAAGGTGA	GGCCCTTCCC	CTTCTGAAGA	960
TCAGGACCTG	GGGATCTTTT	GTGTTGCTCT	TAAGTCCTCC	ACATAGTCCT	GATAGGACTC	1020
CTAAAAAGCA	TTTCAGTGCC	ATCACAAAAC	AAGTAGAGCT	GGGTAGAGCT	GGGCGCGGTG	1080
GCTCACGCCCT	GTAATCCCAG	CACTTTGCGA	GGCCAAGGCG	GGTGGATCAC	GAGGTCAGGA	1140
GTCCAAAACC	AGCCTGGCCA	AGATGGTGAA	ACCCTGTCTC	TACTAAAAAT	GCAAAAAAAT	1200
CAGCCGGATA	TGGTGGCGGG	CGCCTGTAAT	CCCAGGTATT	GGGGAGGCTG	AGGCAGAGAA	1260
TTGCTTGAAC	CCAGGAGGCG	TAGGTTGCAG	TGAGTGGAGA	TCGTGCCTCT	GCAGTCCAGC	1320
CTGGGTGAAA	GAGCGAGACT	CCGTCTCAA	ATGAAAAAAA	AAAAAGAAAA	CAAGTAGAGA	1380
CTGCAAAAAG	GGAACAGTAC	CGGGAATGTT	GGAGAAAAAC	ATACTACAAT	TAAATCCAAC	1440
ACCCCTGTTG	GTCCTGCTAA	ATGACAGGCA	CTGTGGAAGG	TGCTTGGGAC	TCAGATAAAT	1500
AAGACAAAGA	TCTGCCCATG	GAAAGTTCAC	GTCTGGACCA	TAAGGCATTA	GGTTTCATTC	1560
TGAGCTTCCT	AGTGGCCAAG	GCAAAAAGGA	AATAGAATGG	TTTAGACAGC	TCTCATTTGTC	1620
TGATCAAAGG	TGTTGAGGCA	GAGCACTGAG	GAGGGCCTGG	AGATAAAGGG	TGGGCTGGGG	1680
GTCAGATGCA	GTTATCCCTT	TGCCGACCCT	TTGTTCCCTT	TCCTCAGACA	CACCTGCCAG	1740
GGTTTGTGGA	GCAGGCTGAG	GCTCTGAAGG	CCAAGGGAGT	CCAGGTGGTG	GCCTGTCTGA	1800
GTGTTAATGA	TGCCTTTGTG	ACTGGCGAGT	GGGGCCGAGC	CCACAAGGCG	GAAGGCAAGG	1860

TGAGGTGAGG GGCCTGCAGG GAGTCAGGAC CAGGTAGGAT ATTCTTCTTG TGACCTCTAC	1920
TTTCTCTGCA GGTTCGGCTC CTGGCTGATC CCACTGGGGC CTTTGGAAG GTGAGTGTTT	1980
CCCTGACCGC CACAGGGACA TGGCGGTGCG GGGAGCAGTG GGGGCCCTTG GCCTCTTCAA	2040
GGATTTCTGA CACTTTTCTC TGTCTCTTCT TAGGAGACAG ACTTATTACT AGATGATTCTG	2100
CTGGTGTCCA TCTTTGGGAA TCGACGTCTC AAGAGGTAAA AGTGGAGAGT CCTCTGTGGA	2160
GAAAGTCCTC TGTGGGAGAG AGTCCTCTGT GGGAGAGAGT CCTCTGTGGA GAGGGTCCTC	2220
TGTGGGAAGA GTCGTCTGTG GGGGAGATGT GTGGGAGAGA GTCCTGTGTG GGGAGAGTCT	2280
TCTGTAGGGG AGAGTCCTCT GGGGAGAGAG TCCTGTGTGG GGGAGAGTCC TCTGTGGGGA	2340
GAGTCCTCTG TGTGGAGAGA GTCCTGTGTG GTGGTGAGTC CTCTGTGGGG GAGAGTCCTC	2400
TGTGGGGGGA GTCCTCTCTG GAGTTCCTCT GGGCCCCCTGG CTGTTCCTG CCTGTCTCCA	2460
TGCCCAGCCT CCAAGCCCAG GCTGATGCAG CTGGCTGGGC CCCTCTTTCC GGCAGGTTCT	2520
CCATGGTGGT ACAGGATGGC ATAGTGAAGG CCCTGAATGT GGAACCAGAT GGCACAGGCC	2580
TCACCTGCAG CCTGGCACCC AATATCATCT CACAGCTCTG AGGCCCTGGG CCAGATTACT	2640
TCCTCCACCC CTCCCTATCT CACCTGCCCA GCCCTGTGCT GGGGCCCTGC AATTGGAATG	2700
TTGGCCAGAT	2710

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GCCATCCCAG CAGTGGAGGT GTTTG

25

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TTGAACAGCT CTGCCAGGTT CACC

24

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TGGAGGTGTT TGAAGGGGAG CCAG

24

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

CAGGTTTACC TTGTTCCCTG GCTC

24

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GGGTATGGGA CTAGCTGGCG

20

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CTGGCCAACA TTCCAATTGC AG

22

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

ATGTTATGCA ACCCTTTGCG ACAC

24

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GTGTTTGAAG GGGAGCCAGG GAAC

24

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

AGAGACAGGG TTTCACCATC TTGG

24